DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

PRICE DISCOVERY IN FINANCIAL MARKETS: THE CASE OF THE CAPM

Peter Bossaerts

Daniel Kleiman

Charles Plott

SOCIAL SCIENCE WORKING PAPER 1032

April 1999

Price Discovery In Financial Markets: The Case Of The CAPM[1]

Peter Bossaerts, Daniel Kleiman and Charles Plott

California Institute of technology

19 April 1999

[1]Bossaerts and Plott: California Institute of Technology; Kleiman: Lehman Brothers. Bossaerts thanks the Yale School of Management for their hospitality during the period when this draft was written. Preliminary versions of this paper circulated under the title "Experimental Tests Of The CAPM As A Model Of Equilibrium In Financial Markets." The many comments during a seminar at the Yale School of Management, at the Conference on Microstructure and High Frequency Data in Paris (December 1998) and at the Conference on Price Discovery organized by the SFS in Toulouse (March 1999) are gratefully acknowledged. Steve Ross reminded the authors of the opportunities created by the completeness of the markets in the experiments. The financial support of the National Science Foundation and the California Institute of Technology Laboratory for Experimental Research in Economics and Political Science is gratefully acknowledged. Address: m/c 228-77, California Institute of Technology, Pasadena, CA 91125; Phone (626) 395-4067; e-mail: pbs@rioja.caltech.edu, dkleiman@cco.caltech.edu and cplott@hss.caltech.edu.

Price Discovery In Financial Markets:
The Case Of The CAPM

Abstract: We report on experiments of simple, repeated asset markets in two risky securities and one risk-free security, set up to test the Capital Asset Pricing Model (CAPM), which embeds the two most essential principles of modern asset pricing theory, namely, (i) financial markets equilibrate, (ii) in equilibrium, risk premia are solely determined by covariance with aggregate risk. Slow, but steady convergence towards the CAPM is discovered. The convergence process, however, halts before reaching the actual equilibrium. There is ample evidence that subjects gradually move up in mean-variance space, in accordance with the CAPM. Yet, adjustment stops as if the remaining trading time was insufficient to complete all the transactions that are needed to guarantee improvements in positions. We conjecture that this is due to subjects' hesitance in the face of market thinness. Because the convergence process halts, statistical tests reject the CAPM.
JEL Classification: G12, C92, D59. *Keywords*: Capital Asset Pricing Model (CAPM), Experimental Economics, Financial Markets, Equilibrium, Equilibration.

1 Introduction

In many respects, the Capital Asset Pricing Model (CAPM) of Sharpe [1964] and Lintner [1965] is the cornerstone of modern finance. Its decision-theoretic foundation, mean-variance analysis, has become a major guidance to asset allocation. Its equilibrium restriction provides the most important risk correction in the evaluation of portfolio performance. It is widely applied to determine appropriate discount rates in capital budgeting. Asset pricing models with even greater generality are based on CAPM's key arguments of optimal portfolio demands and market equilibration, and share its main prediction, namely, that expected returns increase with the covariation with aggregate risk.

Despite the prestigious theoretical standing of the CAPM (and its more general offspring), there is far less convincing evidence about its empirical relevance. Fama and French [1992], for instance, discovered a significant, negative relationship between average returns and covariation with market returns, in contrast with the prediction in the CAPM of a positive relationship. Frustrated with the productivity of basic model testing, academic finance has ostensibly been drifting towards a purely empirical (statistical) approach to modeling asset prices. E.g., Fama and French [1996], Brennan, Chordia and Subrahmanyam [1996]. The purpose of this branch of research is to measure average returns, instead of to test the underlying theory. The resulting measurements have proven to be very helpful in applied areas, such as capital budgeting and portfolio performance evaluation. Still, the empirical approach is not entirely data driven. Theoretical arguments are often appealed to, mostly to provide a rough indication that the measurements are reasonable. These "reality checks" are necessitated by the limited possibilities of replicating the empirical results on independent samples.

The CAPM rests on very general principles of market behavior which permeate modeling in asset pricing theory, namely that financial markets equilibrate, and that only aggregate risk will be priced in equilibrium. Therefore, it is important to know whether the empirical rejections that are documented in the literature are a mere consequence of some auxiliary assumptions (the maintained hypotheses) that empiricists were forced to make, instead of a reflection of a failure of the general principles behind the model. Several such auxiliary assumptions are dictated by the realities and limitations of empirical work. Some examples are: (i) the empiricist uses a benchmark portfolio as a measure of market risk and return that is not the true market portfolio and perhaps not a sufficiently close proxy; (ii) the empiricist assumes that markets are informationally efficient, in the sense that they hold unbiased beliefs at all times; (iii) the empiricist assumes that markets are continuously in equilibrium, i.e., that observed prices reflect equilibrium; (iv) the empiricist assumes that investors view the world as essentially static, and that their horizon coincides with his/her observation interval. These auxiliary assumptions come on top of the underlying basic principles and are forced upon the empiricist by tractability, computability or mere lack of information.

Empirical rejections amount to rejections of all the assumptions, basic as well as auxiliary. Field work cannot discriminate the causes. The role of experiments, then, is to clarify whether the basics of the CAPM are more than elegant mathematical constructions. An examination of the operations of simple markets can help determine if the underpinnings are scientifically sound, while at the same time revealing some of their critical aspects.

The CAPM is based on very general principles about market behavior and clearly general principles should apply in simple and special cases. If they do not, then they are not general. Experiments, therefore, represent an attempt to go back to basics. By creating real, but simple markets, the operation of the principles can be revealed apart from the auxiliary assumptions used as tools for measurement. In this manner the foundations of the model can be studied and the circumstances under which it might be expected to hold can be isolated with better precision. Knowledge of relationships between model accuracy and circumstances can lead to more successful applications in the more complex field applications where the model is traditionally applied.

Simple markets were created that featured the major properties used by the CAPM model. Only three assets were traded, two risky and one risk-free security. The payoff structure of the risky assets was transparent. Each situation was replicated several times. Markets were organized by the computerized multiple unit double auctions, which is known to successfully facilitate rapid price discovery. Participants were drawn from the Caltech student body. All had some rudimentary training in finance (most were participating as part of an investments class). They were generously paid for good performance (or incurred significant debts for bad performance).[2] Therefore, in a real sense, an attempt was made to give the CAPM the best chance of manifesting itself.

Even in this tightly controlled environment, the CAPM is not a foregone conclusion. The model is based on an equilibrium argument: prices, demands and supplies are revealed simultaneously and coherently, and clearing is instantaneous. In any realistic market situation, the process of equilibration (if it exists at all) is gradual, with feedback effects from temporary prices to demands and supplies. Moreover, the potential profits from speculation about the price discovery process itself (i.e., where prices will move next) may entice some individuals who would otherwise trade up to mean-variance efficient positions. If individuals refrain from mean-variance optimization because they perceive higher profit opportunities from the price discovery process, the CAPM would appear to be a lost cause.

Finally, even if markets equilibrate, it is not clear that equilibrium expected returns will solely depend on covariation with aggregate risk, which is what the CAPM predicts. In other words, one cannot take for granted that the expected returns increase cross-sectionally in the size of the covariation with aggregate risk. In the CAPM, aggregate risk is measured by the return on the market portfolio.

The experiments reported on in this paper were conducted to shed light on these issues. Section 2 presents an overview. Section 3 provides a detailed discussion of the experimental design. Section 4 elaborates on the predictions of the CAPM and more general asset pricing models about the behavior of returns in the experiments. Section 5 explains how standard tests of the CAPM were adjusted to accommodate the unique features of experimental data. Section 6 provides a detailed report and analysis of the experimental results. Section 7 concludes. The first appendix summarizes important technical facts about the CAPM. The second one provides the instruction sheet that was used in the experiments.

2 Overview

The overall results are best captured by the two yardsticks that are traditionally used to measure the performance of the CAPM: the difference between the market's Sharpe ratio and the maximum Sharpe ratio, and the intercept of the Security Market Line (related to Jensen's performance index). Using these measures, we observe a prominent trend towards the CAPM. Convergence is slow, however. Its presence becomes most evident only across sessions. Noise obstructs the trend within experimental sessions. So, there appear to be powerful forces that move the market in the direction of the CAPM, even if they grow only gradually over time.

[2] The best paid subject got away with \$163 for a three-hour experiment; the worst-paid incurred \$55 in debt, which was worked off at \$7 per hour and through the earnings from other experiments.

One should not be surprised about the low speed of convergence. The CAPM is a model about multiple markets with complementary goods (in equilibrium, risky securities have to be held in fixed proportions, no matter how much one decides to hold in total). The theoretical analysis of some price discovery systems (in particular, tatonnement) has long revealed that convergence may be slow or even nonexistent in multiple markets with complementary goods.

Yet, prices stop short of the CAPM equilibrium. That is, prices settle at a point when the market portfolio is still statistically and economically mean-variance inefficient. Because the experiments used a complete set of markets, one could verify that markets did not actually reach another equilibrium, by computing implied state-price probabilities from transaction prices. These reflected an unacceptably high probability of arbitrage opportunities, indicating that markets were still not in equilibrium.

We conjecture that the inability of the market to move all the way to the CAPM equilibrium is caused by market thinness. Participants chose not to rebalance anymore when their portfolios had moved sufficiently close to optimality. This behavior is optimal, because marginal improvements in a portfolio require one to execute trades simultaneously in at least two markets. The chosen market architecture does not allow this. When only part of a desired portfolio rebalancing is achieved, the portfolio may actually have moved to a position that is dominated in mean-variance space by the status quo. Hence, participants sometimes rationally decide not to engage in transactions that would move them all the way to the mean-variance frontier.

Much thought was devoted to the statistical tests of the CAPM. Because the simple experimental markets do not have many of the unknowns of naturally occurring markets, some imagination was required. In empirical studies, the transaction price that is observed at a particular moment in time (e.g., the closing price on the last day of a month) is treated as the actual cost of acquiring a share. All uncertainty is attributed to the end-of-period payoff, whose distribution is unknown. In an experimental setting, the latter is known (being part of the experimental design). Hence, it would be inappropriate to blindly apply the statistical methodology with which one tests the CAPM in field data. In contrast, the market thinness of the experiments implied a substantial uncertainty as to the true cost of acquiring a share. Hence, a statistical methodology was built around the randomness of the acquisition cost, and not the uncertainty about the parameters of the distribution of the final payoff.

The CAPM has rarely been the subject of experimental research. Levy [1997], however, has recently reported results from a CAPM experiment. His main argument in support of the CAPM is based on his observation that the average return/beta relationship is positive, and that volatility does not provide incremental explanatory power of the cross-section of average returns. The present paper is more demanding, by testing the prediction in the CAPM that the average return/beta relationship is *proportional*, or equivalently, that the reward-to- risk ratio (Sharpe ratio) of the market portfolio be the highest possible. At the same time, Levy's environment is simplified substantially, by limiting the number of risky assets to two and the number of possible states to three.

In contrast to Levy, payoffs have nonzero correlation. By using the time-tested continuous double auction, the price uncertainty that is inherent in Levy's batch market system is reduced substantially. The risk-free rate is made endogenous. With only two risky securities, the risk-free rate plays a pivotal role in determining whether the CAPM holds. Experiments were replicated several times, unlike Levy. Finally, the present paper uses a statistical methodology that is better suited for experimental data than standard empirical tests.

3 Experimental Design

The main features of the experiments are summarized in Table 1. A total of seven experiments were conducted. They are indexed by the date of the experiment. Subjects were recruited from the Caltech student community. Many were familiar with basic investment theory, including mean-variance analysis and the CAPM, through an investment class on campus. About one-third of the subjects were graduate students from the natural sciences. The remainder of the subjects was undergraduates, mostly in their junior and senior years. Subjects in the later experiments (5/19 and 6/9) were drawn exclusively from those who participated earlier and were thus experienced with the experimental setting, markets, etc. The number of subjects varied from a low of 5 (4/30 experiment) to a high of 13 (5/13 experiment). With only five subjects, there was insufficient trade in the riskfree security to make reliable inference, as documented below.

While the details of the experiments are outlined in the paragraphs that follow, a brief summary of the setting might be useful. Each experiment consisted of multiple replications (periods) of the same set of conditions. Three securities were created, denoted A, B and C. They had a life of one period, at the end of which they paid a single dividend/payoff and after which they were removed from the system. The magnitude of the dividends depended upon a random draw of one of three possible states, X, Y and Z. The state was drawn after the period was closed, so during trading there was no insider or asymmetric information in the markets. The magnitude of dividends for each of the securities and each state are in Table 1. As can be seen there, the dividend of A varies dramatically with the state, the dividend of B varies much less, and the dividend of C does not vary at all with the state.

All trading took place in an experimental currency called francs, which had a known conversion rate into U.S. dollars (shown in Table 1 for each experiment).[3] Each period traders were endowed with units of A and B (shown in Table 1) and no units of security C. In addition each trader was endowed with a fixed number of francs. It was not possible for holdings of francs, security A or security B to go negative (no short sales) but it was possible for holdings of C to go negative.

[3] In the 4/30 experiment, the experimenter changed the exchange rate *ex post*, from 0.02 to 0.005 dollar per franc. A high number of occurences of state X led him to believe that the (five) subjects would have been paid too generously under the original exchange rate. At the beginning of subsequent experiments, it was explicitly announced that such ex-post changes in the terms of payment would not be made.

Thus, an individual could use francs to buy A and B but the ability to do so was limited by the endowment of francs. If the individual wanted to increase holdings of A and B beyond the implicit holdings of francs it could be done by selling units of C and paying the dividend, which was the same for all states. Thus, a sale of C is like borrowing. The amount of the (known) dividend will be paid by the person who sold the unit of C to the person who bought the unit. Of course, to the buyer the difference between the price paid and the certain dividend is a risk free return since payment is guaranteed by the experimenter. The price of C was determined in the market so the risk-free rate was determined simultaneously with other rates of return.

Table 1 contains the relevant parameters for each experiment. At the beginning of each period, all subjects were endowed with 400 francs and a number of securities A and/or B. Each experiment consisted of three markets and the dividends were identical and public. These are shown in Table 1. The three securities had the same expected value. As can be seen, security A had a higher variance than security B. Security C had no variance at all. The dividends were the same in all experiments, apart from experiment 6/9, when the parameters were changed. Except in the 6/9 experiment, subjects were always endowed with four units of security A and security B and no units of security C (but they could go short in security C). The probability of the states was 1/3 in all experiments.

The parameters for experiment 6/9 differed from the other experiments in two ways. First, the endowments differed. Three groups were formed. Each subject in Group I was given 8 units of security A; each subject in Group II was endowed with 8 units of security B; and the (single) subject in Group III was endowed with the market portfolio (4 of A and 4 of B). Groups I and II had an equal number of subjects. This way, the market portfolio consisted of an equal number of securities A and B, as in all other experiments. Second, the dividends were changed. While the skewness of the A security was positive in the other experiments, it was negative in 6/9. The distribution of security B was left unchanged, as were the expected values. The changes were made in an attempt to increase the liquidity of the markets and to check the robustness of results from the previous experiments.

Each period, earnings were determined as (i) the total payoff (unit dividend times number of units, including the possibility of negative holdings of C) on the securities in inventory at the end of the period, plus (ii) the change in cash position of the period (i.e., end-of-period cash holdings minus the beginning-of-period cash holdings of F400), minus (iii) a pre-determined payment for the endowments given at the beginning of the period, namely 1500 francs. At the end of the experiment, subjects calculated their earnings and were paid in cash. Any subject that lost money was required to work it off. This requirement was especially necessary because of the ability of subjects to borrow (sell C). It was thus possible for a subject to borrow money, buy securities and then suffer losses due to unfortunate draws of the state. In order to prevent the possibility that such bankruptcies would destroy the integrity of the incentive system, the subjects signed an agreement to work off losses.

Subjects were informed that the experiments could last approximately three hours. Once assembled at the Caltech Laboratory for Experimental Economics and Political Science, instructions, included here as an appendix, were read to them. Markets were organized as a computerized, continuous double auction, employing the MUDA technology (Plott[1991]). The experiment started with a trial period in which subjects could familiarize themselves with the technology and procedures by buying and selling with no monetary payoff attached. When the actual experiment started the length of each period was pre-announced and fixed (8 minutes in all experiments, except 4/30, where it was 6 minutes). The number of periods in the entire experiment was not pre-announced but it was determined by the duration of the experiment, from two to three hours.

4 Outcomes Of The Experimental Environment Predicted By Asset Pricing Theory In General And The CAPM In Particular

Modern asset pricing theory was developed with the aim of explaining the empirical properties of returns in financial markets such as the ones described in the previous section. It derives necessary properties for returns using a particular logical argument in which equilibrium plays a prominent role. According to the theory, the nature of financial instruments, financial markets and investor preferences impose special conditions on market equilibrium. The essence can be summarized as follows. Investors demand portfolios that are optimal in a decision-theoretic sense. Prices, and, hence, the distributional properties of returns, adjust as dictated by the laws of market equilibration, so demands adjust to meet a given supply of securities. Therefore, equilibrium prices and returns will reflect the origin of investors' demands, namely, the portfolio-theoretic optimality. This generally implies that average returns on individual securities will increase with covariation with aggregate risk.

While this description of asset pricing theory may leave the impression of a sequential process, all stages are supposed to be *simultaneous and instantaneous*. That is, asset pricing theory focuses on the equilibrium outcome. It is silent about the process of equilibration, and whether the equilibrium can be reached at all with a simple trading mechanism. Asset pricing theory exclusively studies the characteristics of returns that result from the interaction between demand for optimal portfolios and a given supply of securities under the assumption that they are locked in an equilibrium relationship.

The CAPM illustrates this in a simple way. Investors demand portfolios that are optimal in the mean-variance sense, i.e., they minimize variance for a given mean return. At the core of the derivation of the CAPM is the mathematical property that the set of mean-variance efficient portfolios is convex: any weighted average of optimal portfolios is optimal as well. This result facilitates the analysis of equilibrium, as follows. Investors demand mean-variance optimal portfolios. Since the combined (aggregate) demand is a convex linear combination of the individual demands, the aggregate demand is an optimal portfolio as well. In equilibrium, demand must meet supply. Consequently, the portfolio that is supplied to the market must be mean-variance optimal for equilibrium to hold. The supply is usually referred to as the *market portfolio*.

Therefore, the CAPM essentially states that the market portfolio will be mean-variance efficient in equilibrium. Mean-variance efficient portfolios are identified with the property that the expected returns (in excess of the riskfree rate) will be proportional to the covariance with the return on the portfolio. Hence, in the CAPM equilibrium, mean excess returns are proportional to covariance with the return on the market portfolio.

Empirical studies have either focused directly on the mean-variance efficiency of the market portfolio, or on the proportional relationship between average excess returns and covariances that this implies. The first branch of empirical work has essentially tested whether the reward-to-risk ratio of the market portfolio is the highest one can get in the marketplace. The reward-to-risk ratio must be the highest possible, otherwise the market portfolio would not be mean-variance optimal. The reward-to-risk ratio is commonly referred to as the *Sharpe ratio*. Let R_{Ft} denote the return on a risk-free security in period t (we will assume throughout that a risk-free asset exists); let R_{mt} be the return on the market portfolio; let σ_{mt} denote its volatility (standard deviation). The return on the market portfolio is viewed as a random variable and the nature of the parameters on this distribution are part of the auxiliary assumptions imposed by the practical aspects of empirical work. The Sharpe ratio of the market portfolio is defined to be

$$\frac{E[R_{mt} - R_{Ft}]}{\sigma_{mt}}. \tag{1}$$

In practice, the Sharpe ratio of the market portfolio must be estimated, and will generally be less than the maximum Sharpe ratio that was recorded in any given period, because of sampling error. Hence, standard tests of the mean-variance optimality of the market portfolio verify whether the difference between the actual maximum Sharpe ratio and the Sharpe ratio of the market is statistically significant. If it is, one must reject the CAPM. We shall refer to this test as the Sharpe ratio test, SR test. Gibbons, Ross and Shanken [1989] is an example of an empirical study based on the SR test.

Other studies focus on the implication of mean-variance efficiency for the relationship between mean excess returns and covariances with market returns. The latter are usually measured indirectly, by the slope coefficients of projections of security returns onto those of the market portfolio, referred to as "betas."

Specifically, let R_{it} denote the return on security i in period t, and let R_{et} be the return on a mean-variance efficient portfolio. Consider now the (orthogonal) projection of $R_{it} - R_{Ft}$ onto $R_{et} - R_{Ft}$:

$$R_{it} - R_{Ft} = \alpha_i + \beta_i(R_{et} - R_{Ft}) + \epsilon_{it}. \tag{2}$$

The portfolio with return R_{et} is mean-variance efficient if for all i,

$$E[R_{it} - R_{Ft}] = \beta_i E[R_{et} - R_{Ft}].$$

In words: there is a proportional relationship between the mean excess return on any asset i and its "beta." Of course, a security's beta is merely its covariance with the return on the optimal portfolio, divided by the latter's variance. The relationship is proportional, the constant of proportionality being determined by the mean excess return on the mean-variance efficient portfolio.

The CAPM states that the market portfolio is mean-variance efficient. Hence, the CAPM implies the following restriction on equilibrium returns: for all i,

$$E[R_{it} - R_{Ft}] = \beta_i E[R_{mt} - R_{Ft}]. \tag{3}$$

Eqn. (3) could readily be tested by projecting excess returns of individual securities onto excess returns of the market portfolio. If the intercept (α_i in (2)) is statistically significant, the CAPM must be rejected. In fact, this test turns out to be identical to the SR test. See Gibbons, Ross and Shanken [1989]. Instead, tests of (3) have focused on what has become known as the *security market line* (SML), namely, the plot of mean returns (now *not* in excess of the riskfree rate) against market betas. Specifically, the CAPM implies that the SML is (i) one-to-one, and (ii) linear, with (iii) intercept equal to the riskfree rate, and (iv) a positive slope. Early tests following this route include Black, Jensen and Scholes [1972] and Fama and MacBeth [1973]. We will refer to the second test as the SML test.

While the CAPM can be obtained as a model of equilibrium in financial markets only under specific assumptions about payoffs and preferences, more general asset pricing models use analogous arguments and generate qualitatively similar conclusions, with the exception of Ross' [1976] Arbitrage Pricing Theory (APT). Pricing models are based on market equilibration, and invariably imply that expected returns increase with the covariation with aggregate risk. Consequently, if experiments that are meant to induce the CAPM fail to generate evidence in its favor, the relevance of more general asset pricing models must be questioned as well.

Our experiments were designed to provide a favorable environment for the CAPM, i.e., to give it a "best chance." Notice, however, that there were three possible states. Three securities could be traded, with linearly independent payoffs (see Table 1). Consequently, the markets were complete. This has a major advantage, namely, that *all* equilibria can be characterized, whether CAPM or not. Indeed, the complete-market structure implies the existence of a representative agent with expected-utility preferences who supports equilibrium prices. While it is impossible to determine the preferences of this representative agent, in the absence of information on subjects' preferences, prices will not be consistent with the preferences of any representative agent if there are arbitrage opportunities (Harrison and Kreps [1979]). The latter will occur when the (unique) state-price probabilities are not strictly inside the unit interval. Therefore, it is possible to verify that one is at any equilibrium, whether CAPM or not, by checking that the state-price probabilities are all strictly within $(0, 1)$.

State-price probabilities are prices of securities that pay one dollar in one state, and zero in all other, normalized such that they add up to one. State security prices can be implied from the prices of the securities A, B and C by inverting the above payoff matrices.

Why do we care about equilibria that differ from the CAPM? As will be documented in Section 6, prices often did not converge all the way to the CAPM equilibrium. This may reflect one of two phenomena: (i) the market has not converged yet, i.e., the market is still in disequilibrium; (ii) the market has converged to an equilibrium that is close to, yet distinct from, the CAPM equilibrium. By checking the state-price probabilities, these two potential explanations can be distinguished. Nonconvergence to the CAPM can be attributed to the second cause only if all state-price probabilities are reliably within the unit interval. If not, one cannot confidently conclude that any equilibrium was reached.

Theoretically, there are reasons to expect that prices may converge to an equilibrium that differs from the CAPM. If subjects do not have quadratic utility, expected-utility maximization translates into mean-variance optimization only if returns are normally distributed. Therefore, the resulting equilibrium, the CAPM, would obtain only for normally distributed returns. To keep the payoff structure as transparent as possible, however, we chose a three-point distribution. Hence, if one doubts that subjects had quadratic utility, then one would not expect the CAPM to emerge in the first place. From this point of view, it seems that the experimental design did not give the CAPM the best chances to manifest itself. Normally distributed payoffs should have been used.

There is a trade-off, however, between simplicity of a three-point distribution and the best parametrization for the CAPM to obtain. We opted for simplicity, in order not to confuse the subjects. Moreover, the risks from participating in the experiment were not substantial, certainly not career-threatening (nor were they trivial). This means that a quadratic function may provide a good local approximation of subjects' actual preferences. If so, the CAPM would still obtain, despite non-normalities in the payoff distributions.

Security C was risk-free and in zero net supply. Its price was not set, but would be determined by equilibrating demand and supply. Because there was no time value of money, the equilibrium riskfree rate will generally be zero. Often, however, we observed a positive interest rate. That in itself would lead one to reject equilibrium outright. However, it appeared that subjects felt cash constrained, and, hence, started to borrow money in order to execute their buy orders, without waiting for any sell orders to be filled. (As a matter of fact, an understanding of this coordination effort is important to explain the experimental results.) After that, interest rates gradually declined.

5 Determining Statistical And Economic Significance

As mentioned in the Introduction, statistical tests of the CAPM on field data attribute all uncertainty to the final payoff. This is natural, as the parameters of the distribution of the final payoff are unknown. In experiments, however, those parameters are part of the design, and, hence, it would be incoherent to use a statistical methodology that assumes them to be unknown.

Unlike some naturally occurring financial markets, however, trade in our experimental markets was thin. This leads to a substantial amount of transaction price uncertainty. In particular, the closing price can hardly be taken to be representative for the cost of acquiring a share. We take this price uncertainty as an opportunity, however, and build a statistical methodology with which we can coherently test the CAPM on our data.

We proceeded as follows. First, the empirical distribution of the transaction prices was constructed. A weighting scheme was used, because there was often a clear pattern of convergence in transaction prices, so that early observations were less representative of the average cost of a security. Early observations were weighted less heavily. More precisely, if there were T observations in a period, we constructed a new sample with observation t replicated $\max(t, T/2)$ times (T even) or $\max(t, (T+1)/2)$ times (T odd). This generates $\sum_{t=1}^{T} \max(t, T/2)$ (if T even) or $\sum_{t=1}^{T} \max(t, (T+1)/2)$ (if T odd) sample points. The new sample formed the basis for the usual (unweighted) empirical distribution. This particular scheme would generate unbiased estimates even when there is no real trend in transaction prices. (The efficiency would be affected, though.)

Second, the weighted empirical distribution was bootstrapped. The mean of the two hundred bootstraps was taken to be an estimate of the actual cost of a position. From it, an estimate of the expected return was obtained by dividing the expected payoff by the mean bootstrapped transaction price. Likewise, to obtain an estimate of the beta of a security, we divided the covariance between its payoff and that of the market by the bootstrap-based mean transaction price of the security and of the market. The return volatility was estimated by the standard deviation of the payoff divided by the square root of the average transaction price. The riskfree rate was estimated by the payoff on the riskfree asset divided by the mean bootstrapped transaction price.

To test the CAPM, the two approaches from the empirical literature were used. To implement the Sharpe ratio test, the difference between the maximum Sharpe ratio and that of the market portfolio was computed. For the SML test, the SML was estimated from the (estimated) mean returns and betas of the end-of-period positions of all the subjects. Confidence intervals that cover 95% of the outcomes were obtained by bootstrapping the (weighted) empirical distribution, re-deriving the Sharpe ratio test or estimates of the parameters of the SML two hundred times, and computing standard deviations across the outcomes.

Under the null hypothesis that the CAPM holds, the intercept of the SML should be equal to the riskfree rate, and its slope should be positive. The bootstrap-based 95% confidence intervals formed the basis to decide when to reject. One is in a less comfortable situation for the SR test, however. Absent a theory about the distribution of transaction prices in CAPM experiments, one cannot describe the (random) behavior of the average difference between the maximum Sharpe ratio and the market's Sharpe ratio under the null that the CAPM holds. While this may seem annoying, it is also an opportunity to use purely *economic* arguments to construct a decision rule for the SR test.

Let Δ_L denote the left end of the 95% confidence interval of the difference between the maximum Sharpe ratio and that of the market. Δ_H denotes the right end. The decision rule is as follows. *Reject the CAPM* if $\Delta_L > 0.05$. The volatility of the market portfolio was almost invariably close to 0.40.[4] When $\Delta_L > 0.05$, there is more than 97.5% probability that an investor holding the market portfolio could improve her return (at average transaction prices) with 2% ($= 0.05 * 0.40$) or more, keeping volatility constant.

Why is this a reasonable economic benchmark? The 2% benchmark on the improvement in the average return may seem marginal. However, consider subjects who decide to hold the market portfolio. They had to pay F1500 at the end of the experiment for a portfolio with an average payoff of F1600. Hence, their expected income was F100. Their portfolio sold for between F1500 and F1600 in the marketplace, which means that a 200 basis point (2%) improvement in the average return corresponds to an additional income of between F30 and F32, which translates to a 30-32% improvement in income. This is considered to be large enough to reject the CAPM.

Likewise, *do not to reject the CAPM* if $\Delta_L \leq 0.05 < \Delta_H$. Finally, state that there is *strong evidence in favor of the CAPM* if $0.05 \geq \Delta_H$. The latter means: "There is less than 2.5% probability that an investor holding the market portfolio could have improved her average income with 30% or more (at average transaction prices), keeping volatility constant."

6 Results

Presentation and analysis of the experimental results will proceed in four steps. Following the tradition of the experimental literature, a visual impression of the performance of the CAPM in the experiments will be provided first. For that purpose, a plot of the (dynamic) evolution of the difference between the maximal Sharpe ratio and that of the market will be presented. Likewise, the evolution of the intercept of the SML will be plotted against that of the riskfree rate. Second, the results will be evaluated statistically and economically, ignoring any dynamics, as if the market equilibrates instantaneously. Third, the (static) statistical and economic evaluation will be contrasted with the (dynamic) graphical evidence, and discrepancies are interpreted as evidence of disequilibrium. That is, convergence to other equilibria than the CAPM is ruled out. Fourth, its is conjectured that the halted convergence is caused by market thinness. The arguments will be substantiated by an analysis of the evolution of potential improvements in the position of individuals and that of the market portfolio.

[4] This lack of variation in the volatility of the market portfolio contrasts with substantial movements in mean returns. It is not clear why this is.

6.1 Visual Representation Of The Experimental Results

It is standard practice in the experimental finance literature to provide plots of the evolution of the prices of the different securities, in order to gauge evidence of convergence. Figure 1 provides an example. It depicts the evolution of the prices in the 5/13 experiment. As is typical in experimental financial markets, one observes within-period convergence patterns. Across periods, however, there are clear trends (e.g., the price of security B drifts upwards), despite the fact that the environment remained identical.

In contrast with earlier experiments, however, the dynamics of transaction prices cannot be plotted against clearly defined price levels that equilibrium theory predicts. Indeed, the CAPM equilibrium does not translate into specific equilibrium price levels, absent knowledge of subjects' preferences. Worse, even if subjects' preferences are known, there are generally multiple equilibria, each corresponding to particular price levels. See the Appendix for a discussion of equilibrium multiplicity. So, we cannot easily interpret the evidence from plots such as Figure 1.

Therefore, we decided to focus on metrics that unequivocally characterize the CAPM equilibrium, namely, (i) the distance between the maximum Sharpe ratio and the Sharpe ratio of the market portfolio, (ii) the intercept of the SML. These metrics are uniquely defined in the CAPM equilibrium, unaffected by preferences and/or multiplicity of equilibria. Except for sampling error, the former should be zero, and the latter should equal the riskfree rate. In experiments, however, price discovery is a very apparent phenomenon. So, we must not expect initial prices to be such that the Sharpe ratio of the market is economically an insignificant distance below the maximum Sharpe ratio; likewise, we must not expect initial prices to generate a SML with an intercept that is statistically insignificantly different from the riskfree rate. Rather, evidence of convergence must be looked for. The emphasis is on dynamics, i.e., on tendencies.

To obtain the evolution of the two metrics, we proceeded as follows. At the beginning of each period within an experiment, we wait until all three securities traded at least once. Using the most recent transaction prices, we compute the maximum Sharpe ratio, as well as the market's Sharpe ratio. Likewise, we estimate the intercept of the SML from an OLS regression, projecting the expected return of individual positions onto their respective betas. Expected returns and betas are computed on the basis of the most recent transaction prices.[5] Subsequently, Sharpe ratio differences and SML intercepts are re-computed whenever a new transaction takes place. When repeated across periods and experiments, this produces a plot of the evolution of the difference between the maximum Sharpe ratio and the Sharpe ratio of the market, as well as that of the intercept of the SML.

Neither metric could be computed in the few experimental periods that one of the securities was not traded. For instance, during several periods in the 4/30

[5] The procedure differs from the statistical analysis in the next subsection, where we use average (bootstrapped) transaction prices to compute Sharpe ratios, expected returns and betas. This methodology was explained in the previous section.

experiment, the riskfree security (asset C) was not traded. So, no transaction prices were recorded, and, hence, the Sharpe ratios could not be calculated. In that case, we did not estimate the SML either, because its intercept could not be evaluated against a riskfree rate (which would be the return on asset C).[6]

Figure 2 displays the dynamics of the difference between the two Sharpe ratios. Experiments are delineated with vertical lines and labeled with the date when they took place. The horizontal axis measures time in experimental periods. In other words, each period within an experiment covers one unit of time. Zero is the minimum difference between the two Sharpe ratios. Most often, however, the market's Sharpe ratio is below the maximum one, as suggested by the negative Sharpe ratio differences.

Figure 2 exposes a pronounced tendency for the difference between the Sharpe ratios to diminish over time. This is apparent within experiments, but the trend is most manifest across the seven experiments (which were arranged in chronological order). Because of the scale, within- period convergence is less visible. In fact, Sharpe ratio differences often moved erratically within periods, with no clear sign of convergence. Hence, the trend towards the CAPM is most evident over longer time horizons.

With the 5/19 experiment, the market appears to have reached the CAPM equilibrium. To evaluate the robustness of this finding, we altered the payoff structure (reversing the skewness of the payoff of one of the risky securities), as well as individual endowments (allocating only one of the two risky securities to each subject, instead of the market portfolio), as explained in Section 3. Figure 2 clearly suggests that this has no impact on the results: the market's Sharpe ratio is as close to the maximum Sharpe ratio in the 6/9 experiment as it was in the 5/19 experiment.

Figure 3 displays the evolution of the intercept of the SML. To gain perspective, we also plot an estimate of the riskfree rate in each period. See the solid line. This estimate is based on the average transaction prices of security C (the riskfree security). The average is computed from two hundred draws from the (weighted) empirical distribution. In other words, the estimate is identical to the one that will be used in the next subsection, to evaluate the statistical and economic significance of any deviations from the CAPM that we are observing.

While there is evidence that the intercept of the SML converges, the riskfree rate is often not the target. In the 1/27, 4/20, 4/30 and 6/9 experiments, we do observe convergence. While the intercept follows the trend of the riskfree rate in the 5/13 and 5/19 experiments, it remains tenaciously too low. The reverse occurs in the 1/18 experiment. Nevertheless, the range of the intercept contracts gradually, with the last two experiments producing hardly any variation.

In summary:

- The evidence from Figure 2 is unequivocally in support of the CAPM,

[6]If the riskfree security (asset C) did not trade, all subjects could only have had a position in risky securities. Since there are only two risky securities, the market would trivially be on the mean-variance efficient frontier of risky assets only. Hence, the SML would fit perfectly. See the Appendix for further discussion.

whereas that of Figure 3 is mixed. Both figures do demonstrate convergence, but they indicate that it is slow. It takes several experiments for the range of the Sharpe ratio difference and the intercept of the SML to decline.

The slow convergence may suggest that experience is necessary for the CAPM to emerge (remember that the subjects in later experiments had all participated in earlier ones). Still, a peculiarity of the CAPM may have played a role. Indeed, in the CAPM equilibrium, risky securities are like complementary goods (the more one wants of one, the higher the optimal investment in another). Simple tatonnement convergence has long been known to be problematic in the presence of complementary goods.[7] Of course, our experiments were organized as a computerized double auction, as opposed to a standard tatonnement. Nevertheless, earlier experiments with computerized double auctions have produced analogous instabilities (Plott and George [1997]). Which leads one to suspect that complementarity may in part have caused the slow convergence in our CAPM experiments.

6.2 Statistical Tests

Let us now turn to tests of the CAPM along standard empirical methodology. As mentioned in Section 5, however, we do alter the methodology, because a blind application on experimental data would be inappropriate. In particular, we gauge statistical significance using the randomness in transaction prices, as opposed to the randomness in final payoffs. Also, absent a clear model about transaction price uncertainty under the null that the CAPM holds, our decisions to reject or accept the CAPM are based on the economic magnitude of the observed deviations.

Standard empirical methodology assumes, however, that the market is continuously in equilibrium. The plots that we discussed in the previous section clearly indicate the presence of price discovery: while there is strong evidence of convergence, it is implausible to hold on to the empiricist's assumption of continuous equilibrium. Nevertheless, we do want to generate statistics that are directly comparable with the empirical literature. Hence, to mitigate the impact of price discovery, we weigh early transactions in an experimental period less heavily, along the lines of the procedure explained in Section 5. In some experimental periods, we could not perform the tests, because transaction prices for the riskfree security were missing.

Table 2 lists the end-points of the 95% confidence intervals for the difference between the maximum Sharpe ratio and the Sharpe ratio of the market, along with our decision to reject/fail to reject/accept the CAPM (using the rule suggested in Section 5). Only occasionally do we find clear acceptance (when the entire 95% confidence interval is below 0.05). Most often, however, we observe clear rejections. These occur when the entire 95% confidence interval is above

[7] H. Scarf was the first to point this out. A recent reference in this line of research is Hens [1997].

0.05. In economic terms, this would mean that there is a 97.5% probability that average transaction prices are such that an individual holding the market portfolio could improve his/her income by more than 30% (see section 5).

We emphasize the sharp contrast between the statistical evidence obtained from standard empirical tests and the strong visual testimony of convergence towards the CAPM equilibrium from a plot of the evolution of the Sharpe ratio (see Figure 2). When evaluating this discrepancy, one must bear in mind that the statistical tests assume that the market is continuously in equilibrium. Figure 2 clearly proves this assumption wrong. We did try to accommodate an element of price discovery by downweighing early transaction prices in each experimental period, but this appears to be insufficient. Of course, Figure 2 demonstrates that equilibrium discovery usually takes several periods, so that downweighing a few observations within a period won't address the problem adequately.

Table 3 displays the confidence intervals of the intercept of the SML, and contrasts them with the riskfree rate. In addition, 95% confidence intervals for the slope coefficient are provided as well. According to the CAPM, the intercept of the SML should be equal to the riskfree rate, and its slope should be positive. Table 3 suggests that the latter is almost never violated: with two exceptions, the 95% confidence interval is entirely above zero. The CAPM receives far less support from the intercept, however: the riskfree rate is often outside its 95% confidence interval.

Again, the statistical tests in Table 3 assume that the market is continuously in the CAPM equilibrium (although we did weigh early transaction prices less heavily). So, the rejections could be attributed to the methodology's inability to capture price dynamics. In contrast, the plot in Figure 3 captures equilibrium discovery far more aptly. Still, the evidence from Figure 3 matches the statistical results reported in Table 3. For instance, the intercept of the SML stayed below the riskfree rate during the entire 5/13 and 5/19 experiments. Table 3 demonstrates that the distance between the intercept and the riskfree rate over those experiment is significant. On the other hand, Figure 3 may leave one with the impression that the CAPM obtained in the 6/9 experiment. Yet, Table 3 demonstrates that the distance between the intercept of the SML and the riskfree rate was significant during the whole experiment.

In summary:

- The overall evidence from standard statistical tests is discouraging for the CAPM. In part, this should be attributed to the inability of those tests to capture dynamics. Indeed, they assume that the market is continuously in the CAPM equilibrium. Figures 2 and 3 prove that this assumption is wrong. Nevertheless, Figure 3 revealed discrepancies between the intercept of the SML and the riskfree rate even after the market had apparently converged. The statistical tests demonstrate that these discrepancies are significant.

6.3 Summarizing The Evidence And Ruling Out Other Equilibria

Let us combine the evidence from Figures 2 and 3 with the statistical results displayed in Tables 2 and 3. It is fair to conclude that the figures suggest strong, even if slow, tendencies toward the CAPM equilibrium. Yet, at the same time, the process of convergence appears to stop short of the CAPM equilibrium. This explains the statistical rejections reported in Tables 2 and 3. One obviously wonders what keeps the market from moving all the way to the CAPM.

The aborted convergence process suggests that the market remains in disequilibrium. In other words, the market does not equilibrate. Before investigating the causes of this disequilibrium, one does want to rule out the possibility that markets actually converged to another equilibrium than the CAPM. That is, whenever lack of convergence to the CAPM is observed, one must check whether this indeed reflects disequilibrium, because markets may have settled at another (expected-utility) equilibrium. This question is imperative in the 5/13 and 5/19 experiments, as well as the end of the 1/18 and 1/27 experiments: according to the Sharpe ratio differences plotted in Figure 2, the market seems to have converged. The evidence from Figure 3 and Tables 2 and 3, however, indicates that the CAPM has not been reached. Maybe another equilibrium was attained?

In the experiments, a complete set of markets was used. This enables one to reject that prices would not conform to any equilibrium configuration when all state-price probabilities are found to be reliably inside the unit interval. We implied state-price probabilities from the transaction prices of the three traded securities whenever a new transaction occured. In each period, we started computing state-price probabilities only after all three securities traded at least once.

The behavior of the state-price probabilities in each experiment is plotted in Figure 4. The circles indicate the values of the price probabilities of state X (horizontal) and state Y (vertical). For these state-price probabilities to be in the unit interval, and for the third state-price probability to be inside $(0, 1)$ as well, the circles have to be inside the triangles in the plots.

Except for the 6/9 experiment, the state-price probabilities frequently lie outside the triangle, thereby violating the no-arbitrage condition and indicating that the market cannot have been in equilibrium. Take the 5/19 experiment, for instance. Figure 2 suggested that markets had settled, but Figure 3 and Tables 2 and 3 provided ample evidence against the CAPM. Figure 4 demonstrates that the market had not settled at an equilibrium, because state-price probabilities violated the no-arbitrage restriction in almost half the cases.

Figure 5 provides a better view of how the state-price probabilities evolved over time. The price probability for state X is well behaved, fluctuating around 1/3. The price probability for state Y is often fairly high, and that for state Z frequently moves outside the unit interval. In the 5/19, experiment, for instance, the price probability for state Z fluctuates around zero until the last few periods, when it moves reliably inside the unit interval. This implies that markets where out of equilibrium for most of the 5/19 experiment. It is interesting to notice that the Sharpe ratio difference and the intercept of the SML in Figures 2 and 3 evolve correspondingly: both metrics suggest that the markets moved closer to the CAPM equilibrium by the end of the experiment.

Figures 2 and 3 and Tables 2 and 3 revealed that the 6/9 experiment generated most evidence in support of the CAPM. It deserves emphasis that Figures 4 and 5 indicate that this is also the only experiment where one can clearly reject disequilibrium: all state-price probabilities are reliably inside the unit interval.

Table 4 provides 95% confidence intervals for the three state-price probabilities. These were obtained by resampling from the (weighted) empirical distribution of transaction prices. The table confirms the visual evidence from Figures 4 and 5: the state-price probabilities are not always reliably inside the unit interval. The confidence intervals in Table 4 are univariate. That is, they indicate the middle 95% of the mass of the *marginal* (bootstrap) density of each state-price probability, as opposed to the middle 95% mass of the joint density of the three state-price probabilities together. When zero or one is outside the confidence interval, one must conclude that there is less than a 5% probability that transaction prices imply arbitrage opportunities, thereby rejecting the hypothesis that no equilibrium will support the observed prices. Of course, even if the confidence intervals are entirely within the unit interval, markets may still be in disequilibrium. But the fact that we tend to accept the CAPM whenenver this occurs (see, e.g., the 6/9 experiment) proves that one can rule out this possibility.

In summary:

- The apparent arrests in the convergence process towards the CAPM must not be interpreted as evidence that markets actually reached another (expected-utility) equilibrium. When these arrests occur, transaction prices still reveal a high probability of arbitrage opportunities, which is inconsistent with any equilibrium whatsoever.

6.4 Interpreting Halted Convergence

Altogether, we have strong evidence that markets steadily move towards the CAPM, but that the convergence process often stops short of the CAPM equilibrium. One cannot interpret this as evidence that the market reaches other equilibria, because state-price probabilities frequently reveal arbitrage opportunities, indicating that markets must still be in disequilibrium.

One could be tempted to suggest that the relatively low stakes that subjects are exposed to in the experiments lead to satisficing behavior, as opposed to full optimization.[8] In fact, a closer investigation of the costs and benefits that our subjects faced reveals a more fundamental reason for arrests in the convergence process, one that is fully consistent with the foundation of the CAPM, namely,

[8]The stakes were really not that small. As mentioned before, the subject that performed best collected $163 (for an experiment that took about 3 hours); the worst performing subject lost $55.

mean-variance optimization.[9]

With only a limited number of subjects in the experiments, our market mechanism (the double auction) forces participants to actively seek out transactions when they wish to reallocate their portfolios. This involves posting bids (if buying) or asks (if selling) that are aggressive enough to solicit trade. Portfolio reallocations, however, usually require one to simultaneously execute trades in several markets. Because of the thinness of our markets, there is a fair chance that some of the trades may not be executed, despite efforts to make the quotes as attractive as possible. Unfortunately, the resulting portfolio may sometimes be dominated in mean-variance space by the previous allocation. The risk of ending up with an inferior portfolio may induce an optimizing agent not to try to improve his/her position, refraining him/her from moving all the way to the mean-variance efficient frontier.

This can easily be illustrated graphically. For instance, consider position X in Figure 6. The corresponding portfolio consists of a combination of the riskfree security (security C) and the two risky securities (securities A and B). This portfolio is clearly dominated by the riskfree asset, which earns more despite lower (no) risk. So, the holder may want to improve his/her portfolio's performance, moving towards an all riskfree portfolio, by selling the risky securities. If, however, the holder succeeds in selling only holdings of security A, s/he moves *down* in mean-variance space, towards position Y. His/her new position is dominated by the old one. If she only manages to sell security B, the new position (Z in Figure 6) is not dominated in mean-variance space by the old one, but the holder definitely incurs more risk, which s/he may be averse to. Consequently, the holder decides not to engage in any attempt to improve his/her position, even if it is clearly dominated in mean-variance space. It is the transaction risk caused by market thinness that induces a seemingly satisficing attitude.

There is more. If all subjects are content with a marginally inefficient position, one wonders how the position of the market portfolio is affected. Take, for instance, a situation where all individuals move up to positions that are 200 basis points (2%) below the mean-variance efficient frontier, where the distance is measured in terms of forgone mean return for a given volatility (the vertical distance in plots like Figure 6). Would the market also be 200 basis points off the frontier?

Unfortunately, the answer is no. The market portfolio can be far inferior. Technically, this is because convexity is lost off the frontier.[10] Intuitively, the market portfolio may be at an inferior position because it only includes the risky securities. Figure 7 illustrates this. In both Panels in Figure 7, individuals are positioned along a line that is 200 basis points below the mean-variance efficient

[9]Kroll, Levy and Rapaport [1988], however, did report experimental evidence of plain satisficing behavior in a portfolio context: when investors were given the opportunity to simultaneously change the allocations for several assets, they appeared not to move all the way to full optimality.

[10]Roll and Ross [1994] also made this point. Remember that convexity was crucial to derive the CAPM: individuals demand mean-variance efficient portfolios, so, by convexity, the aggregate demand must be mean-variance efficient; in equilibrium, demand must equal supply, and the latter is referred to as the market portfolio.

frontier. The solid portion of the mean-variance frontier of risky securities only (the parabola) traces the possible positions of the market portfolio. It is clear that the market portfolio can be far more than 200 basis points below the actual mean-variance frontier (of all securities, including the riskfree one). In fact, a comparison of Panels A and B of Figure 7 indicates that the range of possible positions of the market portfolio increases as investors are spread out in terms of volatility. Since investors purchase more volatile portfolios as their risk aversion decreases, one could re-state this as follows: the range of possible positions of the market portfolio augments with the range of risk aversion in the marketplace. The Appendix discusses this in more detail.

Of course, our explanation makes sense only if there is indeed a substantial amount of price uncertainty. Figure 8 proves this. It plots the evolution of the prices of the three securities. The plot is generated like Figures 2, 3 and 5: transaction prices were updated after each transaction; if a security did not trade, the previous transaction price was taken; the first observation in a period occurs after each security traded at least once. Figure 8 reveals a high level of transaction price uncertainty. Most of the uncertainty is concentrated in the prices of the two risky securities, A and B.

In conclusion:

- It is perfectly consistent with optimizing behavior that subjects do not trade all the way up to the mean-variance frontier. Transaction price uncertainty caused by market thinness restrains them. This affects the position of the market portfolio, which could now be far below the mean-variance efficient frontier, even far below the position of the average subject.

The detrimental effect of seemingly satisficing individual behavior on the position of the market portfolio does have important repercussions on standard statistical tests of the CAPM. Remember that they test exclusively whether the market portfolio is mean-variance efficient, either by comparing its Sharpe ratio with the maximum one (the SR test), or by checking the properties of the SML (in particular, whether its slope is positive, and whether its intercept equals the riskfree rate). The tests can easily reject, even if the investors follow the prescriptions dictated by the model, namely, mean-variance optimization.

The evidence from Figures 2 and 3 and Tables 2 and 3 support the above description of problems caused by market thinness. On the one hand, we see strong convergence patterns in the Sharpe ratio differences (Figure 2) and the intercept of the SML (Figure 3). This indicates that the forces behind the CAPM (individuals moving up in mean-variance space) are at work. On the other hand, the figures show that the process stops short of the CAPM equilibrium. Tables 2 and 3 corroborate this statistically. In fact, the strength of the rejections of the CAPM could be related to the annoying property that the market portfolio may be way below the mean-variance efficient frontier even if most, if not all individuals are reasonably close to the frontier.

We can illustrate this reasoning in a different way. Figure 9 displays the evolution of the potential gain for each subject in the 5/19 experiment. The potential gain is expressed as the difference in expected return between the subject's actual position and the mean-variance efficient portfolio with the same volatility. In other words, it is the vertical distance between a subject's position and the frontier in plots like Figure 4. Because of its importance in tests of the CAPM, we also plot the evolution of the potential gain for the position taken by the market portfolio. As mentioned before, the potential gain of the market portfolio is not necessarily a simple average of the potential gains of the individual subjects. It can be as bad or even worse as the potential gain of the worst performing subject, i.e., the subject that is farthest off the frontier.

Figure 9 was generated like Figures 2, 3, 5 and 8: after each transaction, the mean- variance frontier is re-computed, and subjects' positions (as well as that of the market portfolio) are re-evaluated. The first observation in each period does not reflect the first transaction; at least one transaction in each security is needed. Unlike in Figures 2, 3, 5 and 8, however, time is measured in number of transactions.

Figure 9 depicts how subjects gradually move up in mean-variance space. There is a clear tendency for most individuals to decrease their potential gain as time progresses. The convergence is not always monotonous. For instance, in periods 1, 3 and 6, subjects' positions become worse before improving. Also, transaction prices sometimes move drastically against a subject's position, leading to a sharp increase in the potential gain. Apparently, these subjects often did not attempt much to reverse the situation. This occurs mainly when there is little time left to trade (the end of the trading period approaches). As mentioned before, this superficial evidence of satisficing behavior masks optimization in the face of transaction uncertainty (market thinness).

The evolution of the potential gain of the market in Figure 9 (the plus signs) explains why the statistical tests in Tables 2 and 3 reject the CAPM every single period in the 5/19 experiment. The market apparently does not follow the best performers. Most often, its potential gain is a simple average of that of the individuals, which means that it is pulled up by the worst performing subjects. Also, convergence of the potential gain of the market is often non-monotonous. Hence, our simple correction of weighting early transactions less does not work effectively to eliminate biases caused by convergence. This adds to the significance of the statistical tests recorded in Tables 2 and 3.

Although a theoretical possibility (as pointed out earlier), we never observed that the market's potential gain was higher than that of the worst performing subject. Nevertheless, individuals with high potential gains often did draw the market's potential gain upwards. A clear example is the last period of the 4/20 experiment. Figure 10 depicts the potential gains for that period directly in mean-variance space, based on average transaction prices (computed on the basis of our bootstrap, like the statistical results in Tables 2 and 3). Individual positions as of the close of that period are indicated with circles; an arrow points to the position of the market portfolio; one individual held the market. Only two subjects held positions below the line through the market portfolio. Half the subjects (four) held almost perfectly efficient portfolios. This is a nice example of how the market portfolio is dragged down by the worst performers.

An investigation of the evolution of potential gains in other experiments confirmed these findings. We can summarize the evidence as follows.

- Plots of individual gains reveal ample evidence that subjects attempt to move up in mean-variance space, in accordance with the CAPM. They stop their quest, however, if there is insufficient time left, and, hence, when they lack opportunities to complete all the necessary trades to improve their portfolio. In this process, the market's position is heavily influenced by the worst performing subjects.

7 Conclusion

It is fair to conclude that market forces pushed prices in the experiments towards the relationship predicted by the CAPM. The convergence process often stopped short of the CAPM, however. One can conjecture that this is caused by market thinness and the associated transaction uncertainty.

Lest the reader would infer otherwise, the experimental results must not be interpreted as evidence demonstrating the validity of the CAPM in naturally occuring markets like the NYSE. Instead, they would merely demonstrate the soundness of the general principles of modern asset pricing theory, namely, that (i) markets equilibrate, and (ii) in equilibrium, expected returns are solely determined by covariation with aggregate risk. These principles are captured by, among other models, the CAPM. The simplicity of CAPM's framework makes it amenable to experimental examination, and that is what this paper reports on.

Demonstration that the general principles of modern finance are sound comes at an opportune moment. This is because evidence on their validity from naturally occuring markets has been mixed at best. Empirical tests of the CAPM are a prime example. Yet, the general principles of asset pricing theory are widely used as a basis of actual financial decision making, in areas such as portfolio management, capital budgeting and performance analysis.

Further experiments could shed light on the role of market thinness. Recently, the technology to conduct economic experiments on a larger scale has become available. Therefore, it has become possible to extend experimentation to much larger groups of subjects. The conjecture is that more liquid markets facilitate (equilibrium) price discovery and its investigation awaits the application of the new technology.

Further experiments should clarify the low speed of convergence. It was emphasized that one cannot expect equilibration to be fast. After all, the adjustments suggested by the CAPM are very difficult to implement. They involve simultaneous trades in different markets. Moreover, absent common knowledge about other subjects' risk aversion, it is never clear a priori at what prices these alternative markets will settle. Feedback effects on the ultimate equilibrium from changes in the endowments during the price discovery process might exacerbate these difficulties. Worse, equilibrium multiplicity is a definite possibility and complementarities between the risky securities may impede market coordination. Of course, experiments could be designed to test the idea that various market instruments facilitate equilibration, such as the introduction of derivative markets on baskets (indices) of risky securities. Index derivatives may play a role beyond the one that is traditionally associated with them (in fact, theory usually assigns no role to derivatives markets whatsoever). Namely, derivatives could be catalysts in the price discovery process. If this is the case, then one might expect a relationship between the presence of derivatives and the accuracy of the CAPM in predicting outcome.

Appendix 1

We collect here the major facts about the CAPM that are of relevance to the analysis of experimental data. Unless explicitly stated otherwise, we assume throughout that individuals demand mean-variance optimal portfolios.

We start with repeating the single main fact and its implications that have guided empirical research.

Fact 1 *In equilibrium, the market portfolio is mean-variance efficient.*

As mentioned before, this result obtains primarily because of the convexity of the set of mean-variance optimal portfolios. The following two implications have inspired empirical research on the CAPM.

Implication 1 *In equilibrium, the market portfolio has the highest possible Sharpe ratio.*

Implication 2 *In equilibrium, the SML is one-to-one and linear; its intercept equals the riskfree rate R_{Ft} and its slope is positive.*

We now state an observation which is due to Nielsen [1988].

Fact 2 *The model can have multiple equilibria.*

This immediately leads to the following implication regarding experimental results.

Implication 3 *In capital market experiments with the same subjects, the same market and payoff configurations, prices may settle at different market Sharpe ratios and imply different SMLs.*

We did see evidence of this only in the 4/30 experiment. According to the SR test in Table 2, mean-variance efficiency of the market portfolio cannot be rejected in periods 1 and 7 of that experiment. Yet, prices settled at substantially different levels. More information can be obtained by contacting the authors.

The next fact appeals to some properties of the mean-variance efficient frontier. To obtain clean statements, we focus on a particular portfolio that plays a prominent role in mean-variance analysis, namely, the minimum second moment portfolio, MSMP. The MSMP minimizes the noncentral second moment, i.e., $E[R_{et}^2]$, instead of the volatility. The MSMP sits on the lower part of the mean-variance frontier. Also, let R_t^- denote the return on the global minimum variance portfolio of only the risky securities.

Fact 3 *If shortsales in the MSMP are allowed, then in equilibrium, $R_{Ft} \leq E[R_t^-]$.*

This can easily be proven by contradiction. If the riskfree rate is above $E[R_t^-]$, then the mean-variance frontier without shortsale restrictions consists of combinations of positive positions in the riskfree security, financed by selling short the MSMP. Since the MSMP takes positive positions in (some or all) risky securities, but all investors prefer to short the MSMP, there is negative demand for (some or all) risky securities. Since risky securities are in positive net supply, the market cannot be in equilibrium.

The following implication turns the above proposition around.

Implication 4 *If $R_{Ft} > E[R_t^-]$ in equilibrium, there must be shortsale restrictions.*

We often observed that $R_{Ft} > E[R_t^-]$ in our experiments. Because there were shortsale constraints on risky securities, this finding does not necessarily contradict equilibrium.

The presence of shortsale constraints on the risky securities in the experiments does not overturn the convexity of the mean-variance efficient frontier (the crucial ingredient to derive the CAPM). However, it does reposition the frontier.

Fact 4 *If there are shortsale restrictions on the risky securities, then the maximum Sharpe ratio decreases as R_{Ft} increases.*

This fact was taken into account in performing the SR tests to be reported below. The result is obvious from a geometric representation of the problem.

In the experimental configuration, there were only two risky securities. Consider in that case the mean-variance efficient frontier composed *only* of risky securities, which we refer to subsequently as the *mean-variance frontier of risky securities*. If there are only two securities, the market portfolio will always be on the frontier of risky securities, whether the CAPM holds or not. This has the following implications regarding the SML.

Fact 5 *If there are only two risky securities, then the SML of portfolios composed of only risky securities is one-to-one, and is linear, with an intercept generally different from R_{Ft}.*

Because the market portfolio is always on the mean-variance frontier of risky securities only, (3) will hold for some choice of riskfree rate R^*_{Ft}. If the market is globally mean-variance efficient, then $R^*_{Ft} = R_{Ft}$.

When plotting the SML on the basis of portfolios that include riskfree positions as well, however, rejections of the CAPM should become obvious.

Fact 6 *If there are only two risky securities and the market portfolio is mean-variance inefficient, then for the SML not to be one-to-one, linear, or not to have a positive slope, portfolios with positions in the riskfree security must be included in the plot.*

The market portfolio is always mean-variance efficient relative to portfolios of risky securities only. Hence, the SML will be one-to-one and linear. If this still holds when the SML is constructed for all portfolios, including those containing riskfree positions, and if its slope is positive, then mean-variance inefficiency would be contradicted.

We now turn to the study of convergence to equilibrium. One fact stands out.

Fact 7 *The CAPM involves complementary goods.*

The CAPM is based on portfolio analysis, which studies the combination of securities to form optimal portfolios. Optimal portfolios involve taking positions in several risky securities. In the CAPM equilibrium, optimal portfolios have a peculiar feature: when scaling up the risk of one's portfolio, all positions have to be increased simultaneously (financed by selling riskfree securities); when decreasing the risk of one's portfolio, the positions in risky securities have to be simultaneously decreased, investing the proceeds in the riskfree security. Hence, risky securities are complementary goods in the CAPM.

Little is known about convergence to equilibrium. Since optimal portfolios consist of complementary securities, however, we can draw some conclusions using the scarce results that are available for convergence to equilibrium for markets with complementary goods.

Implication 5 *Tatonnement convergence is expected to be slow or even non-existent.*

This follows H. Scarf's well-known study of tatonnement convergence.[11] While the experimental markets are organized around the structure of a double auction, as opposed to tatonnement, they have generated similar instabilities. See Plott and George [1992]. Thus, we do take Implication 5 to be suggestive of features that we may observe in double auctions as well. And, indeed, we did notice slow convergence in our experiments (see Figures 2 and 3).

In the CAPM theory, individuals are supposed to buy mean-variance optimal portfolios. In practice, subjects apparently do not adjust to full optimality. This means that subjects may settle for slightly mean- variance sub-optimal positions. Additional gains are possible, in theory. The phenomenon can be observed in almost all experimental markets. Usually, small trading commissions will guarantee tight convergence. In our context, the behavior is optimal, as will be explained later. At this point, one wonders how this affects equilibrium. The market portfolio will obviously be mean-variance inefficient, but one wonders how far below the frontier it may lie. In other words, what is the range of differences between the maximum Sharpe ratio and the Sharpe ratio of the market?

When there are only two risky securities, a simple result can be derived. Assume that investors trade up to positions which are at a fixed distance in terms of mean return (say, 2%) from the efficient frontier, keeping volatility constant.

Fact 8 *Assume investors settle for portfolios that are a fixed (vertical) distance from the mean-variance efficient frontier. As the range of investors' risk aversion increases, the range of potential differences between the maximum Sharpe ratio and the Sharpe ratio of the market increases as well.*

[11]See Hens for a recent attempt to address the problem.

A geometric argument supports the claim. An investor's sub-portfolio of only risky securities is located at the intersection between a line through the riskfree rate and the investor's actual position (including riskfree securities), and the mean-variance frontier of risky securities. As the range of risk aversion increases, the range of preferred volatilities increases, and, hence, the section of the mean- variance frontier of risky securities that individuals invest in enlarges. Since the market portfolio is an aggregate of individual holdings (convex linear combination), it must be somewhere on this section. As the section enlarges, the range of potential differences between the maximum Sharpe ratio and the Sharpe ratio of the market increases as well.

Figure 7 illustrates the argument. In Panel A, the range of risk aversion is low, indicated by the short range of volatilities that investors settle for; the section of the mean- variance frontier of risky securities where the market could be is commensurately small. Panel B displays a case where the range of risk aversion is large. The section of the mean-variance frontier of risky securities where the market could be is correspondingly large.

Notice that the argument does not depend on the number of individuals in the market. Only the range matters. One could refine it if the distribution of risk aversion is known.

Simple graphical examples should convince one that the differences between the maximum Sharpe ratio and the Sharpe ratio of the market can be quite substantial when the range of risk aversion is high. Conversely, when risk aversion does not differ across individuals, everybody will eventually hold the market portfolio, and, hence, the difference in Sharpe ratios is constrained by the fact that individuals trade up to a certain distance below the frontier.

Let us return to the question why individuals are not fully (mean-variance) optimizers. In our experimental context, there is a plausible reason. Marginal improvements in a portfolio require one to execute trades simultaneously in at least two markets, which our market architecture does not allow. For instance, consider the portfolio that is located at the left end of the solid line 2% below the frontier in Panel B of Figure 7. This portfolio is dominated in mean-variance space by the riskfree security. So, the holder may improve her portfolio's performance by selling all holdings of both risky securities. But this is not an easy task: if the holder manages to sell only holdings of security A, (s)he moves *down* in mean-variance space. In other words, the new position is dominated by her old one! If (s)he only sells security B, the new position is not dominated by the old one, but it involves higher risk, which may not be preferred. In this sense, satisfying behavior is in fact optimal in view of the trading uncertainties imposed by our market structure. Hence:

Fact 9 *Inability to simultaneously trade in multiple markets leads fully rational investors to settle for suboptimal portfolios.*

Facts 8 and 9 demonstrate that the forces that move market prices towards the CAPM equilibrium may not be strong enough, even when all investors are mean-variance optimizers. The forces appear to depend on the range of risk aversion across investors and on the degree of synchrony across markets. The latter is, of course intimately related to market thinness.

Appendix 2

Instruction Sheets. Attached to the back of the paper.

References

Black, F. (1972): "Capital Market Equilibrium With Restricted Borrowing," *Journal of Business* 45, 444-454.

Black, F., M. Jensen and M. Scholes (1972): "The Capital Asset Pricing Model: Some Empirical Tests," in: M. Jensen, ed., *Studies in the Theory of Capital Markets*. New York: Praeger.

Brennan, M., T. Chordia and A. Subrahmanyam (1996): "A Re-Examination Of Some Popular Security Return Anomalies," Anderson School of Management, UCLA, working paper 13-96.

Fama, E. and K. French (1992): "The Cross-section of Expected Stock Returns," *Journal of Finance* 47, 427-465.

Fama, E. and K. French (1996): "Multifactor Explanations Of Asset Pricing Anomalies," *Journal of Finance* 51, 55-84.

Fama, E. and J. MacBeth (1973): "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 71, 607-636.

Gibbons, M., S. Ross and J. Shanken (1989): "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121-1152.

Harrison, M. and D. Kreps (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory* 20, 381-408.

Hens, T. (1997): "Stability of tatonnement processes of short period equilibria with rational expectations," *Journal of Mathematical Economics* 28, 41-67.

Kroll, Y., H. Levy and A. Rapaport (1988): "Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model," *American Economic Review* 78, 500-519.

Levy, Haim (1997): "Risk and Return: An Experimental Analysis," *International Economic Review* 38, 119-149.

Lintner, J. (1965): "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics* 47, 13-37.

Nielsen, L. (1988): "Uniqueness of Equilibrium in the Classical Capital Asset Pricing Model," *Journal of Financial and Quantitative Analysis* 23, 329-336.

Plott, C.R. and S. Sunder (1988): "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica* 56, 1085-1118.

Plott, C.R. and G. George (1992): "Marshallian vs. Walrasian Stability in an Experimental Market," *The Economic Journal* 102, 437-460.

Roll, R. and S. Ross (1994): "On The Crosssectional Relationship Between Expected Returns And Betas," *Journal of Finance* 49, 101-121.

Ross, S. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341-360.

Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance* 19, 425-442.

Table 1
Experimental Design Data

Experiment	Subject Category (Number)	Endowments			Payoff Per State									$/F
		A	B	C	A			B			C			
					X	Y	Z	X	Y	Z	X	Y	Z	
1/18	I (12)	4	4	0	475	50	75	125	275	200	200	200	200	0.03
1/27	I (10)	4	4	0	475	50	75	125	275	200	200	200	200	0.02
4/20	I (8)	4	4	0	475	50	75	125	275	200	200	200	200	0.02
4/30	I (5)	4	4	0	475	50	75	125	275	200	200	200	200	0.005
5/13	I (13)	4	4	0	475	50	75	125	275	200	200	200	200	0.03
5/19	I (11)	4	4	0	475	50	75	125	275	200	200	200	200	0.03
6/9	I (4)	8	0	0	275	50	275	125	275	200	200	200	200	0.03
	II (4)	0	8	0	275	50	275	125	275	200	200	200	200	0.03
	III (1)	4	4	0	275	50	275	125	275	200	200	200	200	0.03

Table 2
Sharpe Ratio (SR) Tests Of The CAPM

Experiment and Period		Δ_L	Δ_H	Choice	Experiment and Period		Δ_L	Δ_H	Choice
1/18	1	0.05	0.12	0	4/30	1	0.01	0.07	0
	2	0.01	0.34	0		4	0.04	0.06	0
	3	0.01	0.07	0		5	0.06	0.14	−
	4	0.34	0.48	−		6	0.12	0.48	−
	5	0.22	0.24	−		7	0.01	0.20	0
	6	0.06	0.23	−		8	0.14	0.26	−
	7	0.24	0.27	−	5/13	1	0.00	0.07	0
	8	0.11	0.19	−		2	0.18	0.22	−
	9	0.17	0.20	−		3	0.12	0.17	−
	10	0.07	0.10	−		4	0.10	0.15	−
	11	0.12	0.15	−		5	0.14	0.19	−
1/27	1	0.75	1.26	−		6	0.17	0.20	−
	2	0.06	0.22	−		7	0.15	0.18	−
	3	0.17	0.23	−		8	0.13	0.16	−
	4	0.19	0.27	−	5/19	1	0.13	0.17	−
	5	0.13	0.22	−		2	0.15	0.18	−
	6	0.16	0.56	−		3	0.14	0.16	−
	7	0.19	0.24	−		4	0.14	0.17	−
	8	0.02	0.08	0		5	0.15	0.18	−
	9	0.11	0.22	0		6	0.17	0.18	−
	10	0.00	0.05	+		7	0.12	0.14	−
	11	0.04	0.15	0		8	0.10	0.12	−
	12	0.00	0.06	0		9	0.11	0.14	−
4/20	1	0.03	0.25	0	6/9	1	0.07	0.09	−
	2	0.02	0.12	0		2	0.10	0.11	−
	3	0.15	0.21	−		3	0.05	0.09	0
	4	0.00	0.02	+		4	0.03	0.09	0
	5	0.00	0.04	+		5	0.03	0.08	0
	6	0.35	0.38	−		6	0.02	0.06	0
	7	0.08	0.28	−		7	0.00	0.02	+
						8	0.06	0.10	−

Remarks: Choice: − if reject CAPM ($\Delta_L > 0.05$), 0 if fail to reject CAPM ($\Delta_L \leq 0.05 < \Delta_H$), + if strong evidence in support of CAPM ($\Delta_H \leq 0.05$). Δ_L and Δ_H are the lower bound and upper bound of the 95% confidence interval of the difference between the maximum Sharpe ratio and the Sharpe ratio of the market, respectively. Confidence intervals are estimated by bootstrapping from the (weighted) empirical distribution of transaction prices.

Table 3
Security Market Line (SML) Tests Of The CAPM

Experiment and Period		R_{Ft}	Intercept		Slope		Experiment and Period		R_{Ft}	Intercept		Slope	
			Low	High	Low	High				Low	High	Low	High
1/18	1	0.2	1.3	2.2	1.4	1.8	4/30	1	4.0	1.4	2.5	0.9	2.7
	2	4.3	7.8	9.2	5.6	6.6		4	4.5	1.9	2.8	1.2	3.3
	3	8.3	6.7	8.1	3.3	5.0		5	5.0	-2.3	1.5	6.3	7.3
	4	4.2	10.5	12.0	16.2	18.9		6	10.0	-1.3	-0.3	7.7	8.6
	5	0.7	6.2	6.6	11.1	11.8		7	6.7	1.4	6.6	2.2	7.2
	6	1.9	5.6	6.2	8.6	9.1		8	9.0	-2.0	0.4	8.1	9.8
	7	0.6	6.2	6.7	17.4	18.9	5/13	1	6.2	3.4	5.1	4.9	5.9
	8	1.2	5.2	6.1	20.2	20.7		2	9.8	1.1	2.6	2.6	3.4
	9	1.0	5.2	5.8	19.7	20.7		3	7.5	0.2	1.7	4.9	6.6
	10	2.0	5.2	6.0	14.6	15.9		4	8.7	2.4	4.3	8.1	10.0
	11	0.5	4.2	5.0	11.3	12.4		5	6.3	-2.4	-0.5	7.2	8.1
1/27	1	1.6	13.5	22.6	-0.3	3.2		6	7.0	-1.6	-0.5	7.1	7.6
	2	1.9	3.3	5.5	2.4	3.2		7	5.6	-3.7	-2.6	8.0	9.2
	3	1.5	5.3	6.0	11.8	12.7		8	5.4	-1.2	0.0	8.5	9.0
	4	1.1	6.8	8.3	18.4	20.1	5/19	1	5.8	-1.0	0.0	3.6	4.1
	5	2.2	6.6	7.8	18.5	19.3		2	4.1	-2.3	-1.9	5.0	5.5
	6	1.1	6.1	12.3	21.7	22.6		3	4.5	-3.7	-3.4	9.1	9.5
	7	1.7	6.7	8.1	12.3	14.9		4	3.3	-4.7	-3.7	5.3	7.4
	8	2.3	3.5	4.3	3.6	4.8		5	3.7	-5.0	-4.2	6.5	7.7
	9	2.1	6.0	8.2	15.3	16.8		6	4.2	-5.8	-5.4	7.9	8.2
	10	4.5	4.0	5.0	4.2	5.7		7	2.8	-4.5	-4.1	6.6	7.3
	11	2.1	3.3	4.7	0.5	1.3		8	3.7	-3.3	-1.3	8.0	8.4
	12	3.0	1.3	2.1	-0.8	-0.3		9	4.7	-1.8	-1.1	7.6	8.3
4/20	1	4.5	7.1	11.8	16.2	23.5	6/9	1	2.5	0.8	1.2	0.4	0.8
	2	7.1	2.7	5.2	2.3	4.9		2	2.5	0.4	0.7	0.4	0.6
	3	9.4	1.9	2.7	0.8	3.0		3	1.9	0.4	0.8	0.3	0.6
	4	4.6	4.3	5.0	2.5	3.4		4	1.9	0.4	0.8	0.7	0.9
	5	5.4	4.3	5.4	2.9	4.1		5	2.0	0.9	1.2	0.2	0.4
	6	0.7	5.7	6.0	5.5	6.2		6	1.6	0.5	0.8	0.4	0.6
	7	5.2	8.0	10.9	4.7	5.9		7	0.9	0.6	0.7	0.4	0.5
								8	2.0	0.0	0.5	0.2	0.4

Remarks: The riskfree rate (R_{Ft}) is an average based on two hundred bootstraps from the (weighted) empirical distribution of transaction prices. “Low” and “High” are the lower bound and upper bound of the 95% confidence interval of the intercept and slope of the OLS fit of the Security Market Line (SML), respectively. Confidence intervals are estimated by bootstrapping from the (weighted) empirical distribution of transaction prices. The riskfree rate, intercept and slope of the SML are all expressed in percentage points.

Table 4
State-Price Probabilities: 95% Confidence Intervals

Experiment and Period		State X Low	State X High	State Y Low	State Y High	State Z Low	State Z High	Experiment and Period		State X Low	State X High	State Y Low	State Y High	State Z Low	State Z High
1/18	1	0.30	0.31	0.26	0.30	0.39	0.43	4/30	1	0.31	0.34	0.37	0.46	0.21	0.32
	2	0.22	0.26	0.15	0.32	0.41	0.62		4	0.30	0.33	0.43	0.46	0.21	0.27
	3	0.28	0.30	0.39	0.50	0.21	0.33		5	0.28	0.29	0.49	0.64	0.07	0.23
	4	0.13	0.15	0.05	0.12	0.74	0.82		6	0.27	0.39	0.62	1.30	-0.69	0.12
	5	0.17	0.18	0.18	0.20	0.63	0.65		7	0.24	0.30	0.39	0.72	-0.02	0.40
	6	0.20	0.22	0.19	0.30	0.48	0.61		8	0.26	0.30	0.67	0.91	-0.22	0.07
	7	0.11	0.13	0.16	0.18	0.69	0.73	5/13	1	0.27	0.30	0.36	0.51	0.20	0.37
	8	0.10	0.11	0.22	0.27	0.63	0.68		2	0.33	0.35	0.70	0.80	-0.14	-0.04
	9	0.10	0.11	0.20	0.23	0.66	0.69		3	0.29	0.32	0.61	0.70	-0.01	0.10
	10	0.14	0.16	0.27	0.29	0.55	0.59		4	0.24	0.27	0.57	0.66	0.07	0.18
	11	0.17	0.18	0.24	0.26	0.56	0.59		5	0.27	0.29	0.66	0.76	-0.05	0.06
1/27	1	0.21	0.24	-0.32	-0.04	0.80	1.10		6	0.28	0.29	0.72	0.78	-0.07	0.00
	2	0.28	0.29	0.21	0.29	0.42	0.51		7	0.27	0.29	0.67	0.73	-0.02	0.06
	3	0.18	0.19	0.19	0.22	0.60	0.63		8	0.25	0.26	0.64	0.72	0.03	0.11
	4	0.10	0.12	0.16	0.21	0.67	0.74	5/19	1	0.31	0.33	0.61	0.68	-0.01	0.08
	5	0.11	0.12	0.19	0.25	0.63	0.70		2	0.29	0.30	0.67	0.71	-0.01	0.04
	6	0.08	0.09	0.01	0.23	0.68	0.91		3	0.26	0.27	0.65	0.71	0.02	0.09
	7	0.15	0.17	0.18	0.22	0.61	0.67		4	0.28	0.32	0.65	0.70	-0.02	0.07
	8	0.26	0.28	0.28	0.32	0.40	0.46		5	0.29	0.30	0.67	0.72	-0.02	0.04
	9	0.13	0.15	0.19	0.27	0.59	0.67		6	0.29	0.29	0.70	0.73	-0.03	0.01
	10	0.26	0.29	0.30	0.43	0.27	0.44		7	0.27	0.29	0.61	0.65	0.07	0.12
	11	0.30	0.32	0.25	0.30	0.38	0.45		8	0.26	0.27	0.57	0.62	0.11	0.17
	12	0.34	0.36	0.32	0.44	0.20	0.34		9	0.27	0.28	0.60	0.64	0.08	0.13
4/20	1	0.08	0.14	0.19	0.33	0.54	0.73	6/9	1	0.26	0.27	0.33	0.34	0.39	0.41
	2	0.29	0.33	0.42	0.59	0.08	0.29		2	0.24	0.26	0.33	0.33	0.41	0.43
	3	0.31	0.34	0.68	0.77	-0.11	0.01		3	0.26	0.29	0.33	0.34	0.37	0.41
	4	0.29	0.31	0.32	0.39	0.30	0.39		4	0.26	0.30	0.33	0.35	0.35	0.40
	5	0.28	0.31	0.33	0.43	0.26	0.39		5	0.28	0.31	0.33	0.34	0.35	0.40
	6	0.24	0.25	0.12	0.14	0.61	0.64		6	0.28	0.32	0.33	0.34	0.34	0.39
	7	0.24	0.26	0.17	0.28	0.47	0.59		7	0.32	0.33	0.34	0.34	0.32	0.34
									8	0.26	0.28	0.32	0.33	0.39	0.42

Remarks: State-Price Probabilities are inverted from transaction prices. Confidence intervals are estimated by bootstrapping two hundred times from the (weighted) empirical distribution of transaction prices.

Figure 1: Evolution of the prices of the three securities (A, B and C) in the 5/13 experiment. Time is measured in periods. Vertical bars delineate periods. Prices are updated whenever there is a new transaction. The first observation in a period is computed only after each security traded at least once.

Figure 2: Evolution of the difference between the maximum Sharpe ratio and the Sharpe ratio of the market portfolio. The theoretical maximum difference is zero. Time is measured in periods. The date of an experiment is indicated between the vertical bars that delineate it. Sharpe rate differences are updated whenever there is a new transaction. The first Sharpe ratio difference in a period is computed only after each security traded at least once. Periods when one of the securities did not trade are left empty.

Figure 3: Evolution of the intercept of the security market line (SML) (asterisks) and the riskfree rate (solid lines). In the CAPM equilibrium, these should be equal. Time is measured in periods. The date of an experiment is indicated between the vertical bars that delineate it. Intercepts are updated whenever there is a new transaction. The first intercept in a period is computed only after each security traded at least once. Periods when one of the securities did not trade are left empty. The riskfree rate is estimated on the basis of 200 bootstraps from the empirical distribution of the transaction prices of the riskfree security in each period.

Figure 4: Evolution of the state-price probabilities of state X (horizontal axis) and state Y (vertical axis) in each experiment. In equilibrium, these state-price probabilities should lie strictly inside the triangle. State-price probabilities are implied from the transaction prices of securities A, B and C, and are re-computed whenever there is a new transaction.

Figure 5: Evolution of the state-price probabilities for the three states (X, Y and Z). In equilibrium, all state-price probabilities should lie strictly inside the interval $[0, 1]$ (betweenhorizontal dotted lines). Time is measured in periods. The date of an experiment is indicated between the vertical bars that delineate it. State-price probabilities are implied from the transaction prices of securities A, B and C, and are re-computed whenever there is a new transaction. The first state-price probability in a period is computed only after each security traded at least once. Periods when one of the securities did not trade are left empty.

Figure 6: An investor who holds position X may want to move to point C, because it has a higher return and lower (no) risk. Since X consists of a combination of the positions A, B and C, both A and B have to be liquidated and additional units of C have to be bought. If the investor manages only to sell A, (s)he moves down to position Y, which is dominated by the original position (X). If (s)he can only sell B, (s)he moves to Z, which has higher expected return, yet is more volatile.

Figure 7: Graphical illustration of the effect of the range of risk aversion (low in Panel A; high in Panel B) on the possible position of the market portfolio when all individuals locate on a line 200 basis points (2%) below the mean-variance frontier. Individuals are positioned on the solid straight line below the (global) mean-variance efficient frontier. The market portfolio will be located anywhere on the solid section of the mean-variance frontier of the two risky securities only.

Figure 8: Evolution of the prices of the three securities (A, B and C). Time is measured in periods. The date of an experiment is indicated between the vertical bars that delineate it. Prices are updated whenever there is a new transaction. The first observation in a period is computed only after each security traded at least once. Periods when one of the securities did not trade are left empty.

Figure 9: Evolution of potential gains on individual positions (dots) and the market portfolio (pluses) in the 5/19 experiment. Potential gains are defined as the increase in expected return one obtains from moving up to the mean-variance frontier, keeping volatility constant. Time is measured in transactions. Vertical bars delineate periods. Potential gains are updated whenever there is a new transaction. The first potential gains in a period are computed only after each security traded at least once.

Figure 10: Portfolio space in period 7 of the 4/20 experiment, based on average transaction prices that were bootstrapped from the (weighted) empirical distribution. Circles depict the positions of the 8 subjects at the end of the period. One subject held the market portfolio. Only two subjects had a higher potential gain (vertical distance to the mean- variance efficient frontier) than the market portfolio.